Physically-informed kernels for wave loading prediction

D.J. Pitchforth[1], T.J. Rogers[1], U.T. Tygesen[2], E.J. Cross[1]

[1] Dynamics Research Group, Department of Mechanical Engineering, University of Sheffield, Mappin Street, Sheffield S1 3JD, UK

[2] Ramboll Energy, Global Division-Jackets, Bavnehøjvej 5, 6700 Esbjerg, Denmark

email: djpitchforth1@sheffield.ac.uk

ABSTRACT:
Wave loading is a primary cause of fatigue within offshore structures and its quantification presents a challenging and important subtask within the SHM framework. The accurate representation of physics in such environments is difficult, however, driving the development of data-driven techniques in recent years. Within many industrial applications, empirical laws remain the preferred method of wave loading prediction due to their low computational cost and ease of implementation. This paper aims to develop an approach that combines data-driven Gaussian process models with physical empirical solutions for wave loading, including Morison's Equation. The aim here is to incorporate physics directly in to the covariance function (kernel) of the Gaussian process, enforcing derived behaviours whilst still allowing enough flexibility to account for phenomena such as vortex shedding, which may not be represented within the empirical laws. The combined approach has a number of advantages including improved performance over either component used independently and interpretable hyperparameters.

KEY WORDS: Offshore structures; Gaussian process; Physically-informed kernel design.

1 INTRODUCTION

The demand for accurate estimation of the lifespan of offshore structures is growing, being driven by both the continued development and construction of offshore renewables [1] and ageing of existing structures. Reducing the risk associated with the large investment required to construct new facilities [2] and the need to assess structural health due many structures nearing the end of their initial 20-25 year design lives [3] are key factors.

The modelling of physics within extreme environments, such as offshore, is challenging and the development of physics-based models is therefore difficult. Phenomena within fluid mechanics including vortex shedding and turbulence may be computationally expensive to model and require expert knowledge to validate. In structural dynamics, the effects of changes in manufacturing tolerances and mechanical joints may induce large changes in dynamic behaviour.

The increased availability of monitoring data for engineering structures has lead to the surge in adoption of data-based methods in recent years, allowing for the direct learning of relationships from data, without the need for complete knowledge of the underlying physical process. For a regression task, such as wave loading prediction, neural networks and Gaussian processes have already shown to be effective in a wide range of structural dynamics applications [4–6]. Although effective when employed correctly, machine learning techniques are not without limitations: extrapolation outside observed conditions often has significant impact on performance and the overfitting of models, particularly highly flexible ones, during training is a problem [7].

A model combining physics and data-based approaches, sometimes referred to as a 'grey-box' [8–11] or 'hybrid' [12–14] model, aims to extract the benefits of both: structure, insight and extrapolative capabilities from physics; and flexibility and the capability to model unknown processes from a data-based component. The construction of grey-box models could be considered to lie across a sliding scale. At one end, one would begin with a physics-based model and add flexibility through the introduction of a data-based component; at the other, one might apply physical constraints to existing machine learning algorithms [15].

The models developed within this paper aim to combine physics-based methods of empirical wave loading prediction with data-based Gaussian process NARX models. There are a wide range of potential means to combine physics and data, with the differences in model structure and performance produced by different methods of combination presenting an interesting research area. Previous work by the authors [11] focussed on residual modelling and input augmentation within a wave loading context. Both methods were found to improve overall predictive performance, with residual modelling offering particular benefit in extrapolation. Far away from previously observed conditions, model predictions revert to the physics-based mean function. Although an effective means of assisting extrapolation where the physics used may be well validated, this form of model structure places a heavy reliance on the mean function used and care should be taken to ensure that any assumptions used in its construction remain valid.

Here, the physics is embedded within the design of the covariance function (kernel). This has a number of potential advantages over a physics-based mean function, including single step learning within a standard GP framework, maintaining of the original signal to noise ratio and ability for separate components in the kernel to account for different contributing factors in a process. The phenomena of mis-learning parts of a process (i.e. components of a target function) when using a prior mean that only reflects some of the behaviour of interest is highlighted in Figure 1. Here a toy function is learned with a GP with a prior mean that only captures the linear component of the process. Although the overall combined fit of the model is good, the capture of the function components by the linear mean function and residual GP does not reflect the process, which may cause problems where training data are few or in an extrapolation task. Additionally, when physics-based models attempt to capture processes excluded by the assumptions present within their construction, a biasing of model parameters can often occur. By jointly optimising the physics and data-based components of the model the misrepresentation of signal components is less likely.

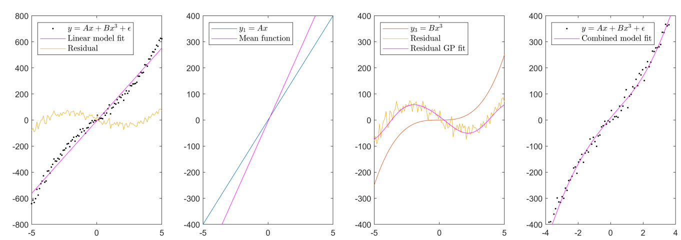

Figure 1: Subplots highlighting the residual modelling process of a synthetic function, $y = Ax + Bx^3 + \epsilon$, and potential misrepresentation of individual function components. From left to right: The fitting of a linear mean function and residual calculation, the modelled function contributions compared with the underlying linear and cubic processes and the combined model fit.

The design of the covariance function here will rely on combining physically-derived kernels with their more traditional data-driven counterparts. Models are implemented on a dataset collected from the Christchurch Bay Tower (CBT) [16], an offshore test facility providing measurements of a real sea state environment. Training and validation occurred on a 500 point subset of the complete dataset with the aim to achieve a comparison between the performance of different model structures rather than maximising performance of a selected final model.

2 EMPIRICAL WAVE LOADING PREDICTION

Empirical methods of wave loading prediction offer a balance between predictive performance, computational resource requirements and ease of model validation; they are popular within many industrial applications [17, 18]. For the modelling of wave loads on slender members, which many offshore structures are comprised of, Morison's Equation has been the most widely used such method since its introduction in 1950 [19].

For a stationary, rigid, slender, cylinder of diameter D positioned within waves of velocity U and acceleration $\dot{U}$, the force per unit axial length F is expressed:

$$F = \underbrace{\frac{1}{2}\rho D C_d}_{C'_d} U|U| + \underbrace{\frac{1}{4}\pi\rho D^2 C_m}_{C'_m} \dot{U} \tag{1}$$

where ρ is the fluid density, C_d is the drag coefficient and C_m is the inertia coefficient. The dimension specific terms may be grouped to form two constants C'_d and C'_m relating to the drag and inertia forces of the wave.

An important consideration when using empirical methods, as with any engineering model, is understanding the limitations and assumptions made within the construction of the model. To achieve their computational efficiency, empirical methods often rely on strong simplifying assumptions. For Morison's equation these include unidirectional flow, the waves being unaffected by the presence of the structure (Water depth $\gg D$) and the separation of force in to drag and inertia components [18–20].

Morison's Equation is generally well regarded within the literature [20–22] with Sarpkaya stating "it is unlikely that an entirely new equation will ever replace it" [23]. Research efforts focus mainly on the development of extensions and modifications to Morison's Equation, rather than a competing alternative. Such modifications include adapting Morison's Equation to work on inclined cylinders [24–26], reducing the number of required model coefficients [27], improving the model fit in cases of sinusoidal flow [28] and improved wave force classification [29, 30].

3 GAUSSIAN PROCESS NARX MODELS

A Nonlinear AutoRegressive model with eXogeneous inputs (NARX) passes previous signal values y_{t-i} and additional (exogenous) inputs u_{t-j} through some nonlinear function $f(x)$.

$$y_t = f([u_t,\ u_{t-1}, ...,\ u_{t-l_u},\ y_{t-1},\ y_{t-2}, ...,\ y_{t-l_y}]) + \varepsilon \tag{2}$$

where l_u and l_y are the maximum lagged time steps considered for the exogeneous inputs and previous signal values; they should be considered as additional model hyperparameters and determined via an appropriate lag selection process [31, 32].

In the case of a GP-NARX, $f(x)$ is a Gaussian process (GP), offering several advantages over alternative NARX model variants. A GP is non parametric and flexible, with no fixed functional form; this allows for the capture of complex relationships within data without extensive prior knowledge of the underlying physical process. Being a probabilistic technique, a GP may also provide quantification of uncertainty alongside predictions. The reader is encouraged to consult [33] for an overview of GP theory.

Multiple types of prediction may be generated from a NARX model, defined by the nature of the lagged target signal values used as the input to the model. For One Step Ahead (OSA) predictions,

$\boldsymbol{y_{t-1:t-l_y}}$ are measured values of the target signal whereas for the Model Predicted Output (MPO), $\boldsymbol{y_{t-1:t-l_y}}$ are previous predictions of the target signal. An extension to the MPO, the Monte Carlo sampled Model Predicted Output (MC MPO) feeds back samples from the full predictive distribution rather than just the mean prediction and has been shown to provide a more realistic capture of model uncertainty [34, 35].

For the operation of models developed within this paper, it is assumed that the measured wave force will be unavailable and therefore the MPO and MC MPO will be the prediction types of interest.

4 MORISON'S EQUATION IN KERNEL FORM

Within a GP, the kernel defines the family of functions from which predictive samples may be drawn. Through the design and selection of kernels, one may control and restrict the behaviour of predictions generated from a GP and enforce desirable or physically derived constraints. To incorporate Morison's Equation (1) within a kernel, it is first appropriate to setup a Bayesian Linear regression.

$$X = [\boldsymbol{U}|\boldsymbol{U}|,\ \dot{\boldsymbol{U}}] \tag{3}$$

$$\boldsymbol{\beta} = [C'_d,\ C'_m]^T \tag{4}$$

$$p(\boldsymbol{F}|X, \boldsymbol{\beta}, \sigma_n^2) \sim \mathcal{N}(X\boldsymbol{\beta}, \sigma_n^2\mathbb{I}) \tag{5}$$

The covariance of a process $f(\boldsymbol{x}) = \boldsymbol{x}\boldsymbol{\beta}$ with prior $\beta \sim \mathcal{N}(0, \sigma_\beta^2\mathbb{I})$ between two input vectors $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$ is derived

$$\begin{aligned} cov(f(\boldsymbol{x_i}), f(\boldsymbol{x_j})) &= \mathbb{E}[f(\boldsymbol{x_i})f(\boldsymbol{x_j})] - \mathbb{E}[f(\boldsymbol{x_i})]\mathbb{E}[f(\boldsymbol{x_j})] \\ &= \mathbb{E}[(\boldsymbol{x_i}\boldsymbol{\beta})(\boldsymbol{x_j}\boldsymbol{\beta})^T] - \cancelto{0}{\mathbb{E}[\boldsymbol{x_i}\boldsymbol{\beta}]\mathbb{E}[\boldsymbol{x_j}\boldsymbol{\beta}]} \\ &= \boldsymbol{x_i}\mathbb{E}[\boldsymbol{\beta}\boldsymbol{\beta}^T]\boldsymbol{x_j}^T \\ &= \boldsymbol{x_i}\Sigma_\beta \boldsymbol{x_j^T} \end{aligned} \tag{6}$$

The expression for covariance can then be used as a kernel within a GP

$$K_{Lin}(X, X') = X\Sigma_\beta X'^T + \sigma_n^2\delta_{X,X'} \tag{7}$$

The use of this linear kernel over input space $X = [\boldsymbol{U}|\boldsymbol{U}|,\ \dot{\boldsymbol{U}}]$ will be equivalent to a Bayesian Linear Regression of Morison's equation, however the training time is now of order $O(n^3)$ rather than $O(n)$. The advantage of kernel representation comes from the ability to use multiple kernels in combination. If one considers the target to be the sum of $f(\boldsymbol{x})$ and some function $g(\boldsymbol{x})$ to account for unmodeled phenomena, then it follows that the covariance structure is

$$\begin{aligned} K(X, X') &= K_{Lin}(X_1, X'_1) + K_{SE}(X_2, X'_2) + \sigma_n^2\delta_{X,X'} \\ &= \underbrace{X_1\Sigma_\beta X_1'^T}_{\text{Morison's Equation}} \\ &\quad + \underbrace{\sigma_f^2 \exp\left(-\frac{1}{2}(X_2 - X'_2)\Lambda^{-1}(X_2 - X'_2)^T\right)}_{\text{Excluded phenomena}} \\ &\quad + \underbrace{\sigma_n^2\delta_{X,X'}}_{\text{Noise}} \end{aligned} \tag{8}$$

where a Squared Exponential (SE) has been used as the covariance of the unknown behaviour under a zero mean assumption. Here $X_1 = [\boldsymbol{U}|\boldsymbol{U}|,\ \dot{\boldsymbol{U}}]$, $X_2 = [\boldsymbol{U_t},\ \dot{\boldsymbol{U}}_{\boldsymbol{t}},\ \boldsymbol{U_{t-1}},\ \dot{\boldsymbol{U}}_{\boldsymbol{t-1}},\ \boldsymbol{y_{t-1}},\ \boldsymbol{y_{t-2}},\ \boldsymbol{y_{t-3}}]$, σ_f^2 is the signal variance, Λ is the matrix of length scales such that $diag(\Lambda) = [l_1^2, l_2^2, ..., l_D^2]$ for a D dimensional input and σ_n^2 is the noise variance. The selected maximum lags, $l_u = 1$ and $l_y = 3$, have been carried forward from previous work [11].

5 CASE STUDY: THE CHRISTCHURCH BAY TOWER DATASET

To obtain a measure of performance within a real sea state environment, models were implemented using a dataset collected from the Christchurch Bay Tower (CBT) [16]. Constructed as a test facility, the structure is equipped with a dense array of sensors including Perforated ball Velocity Meters (PVMs), pressure transducers and crucially, force sleeves, allowing for valuable measurement of the target wave force.

Models were trained and validated on subsets of 500 data points, with performance measured on unseen test set of 1000 data points. All data was selected to maximise the ratio of x-velocity to y-velocity, where the flow was primarily unidirectional.

5.1 Model Predictive Performance

The performance of models was measured using two metrics: the Normalised Mean Square Error (NMSE) and the Mean Standardised Log Loss (MSLL). The MSLL is a probabilistic measure, with superior models having more negative scores. A baseline of zero is equivalent to setting the predictive mean and variance for all test set points as the mean and variance of the training set. A comparison of results is shown in Table 1.

Table 1: Model performance comparison.

Model	Structure	NMSE (%)	MSLL
Morison's Eq.	$X_1\boldsymbol{\beta}$	19.528	-0.813
	$X_1\boldsymbol{\beta}$ *(Normalised)*	15.458	-0.857
GP-NARX MC MPO	$\mathcal{GP}(0, K_{SE})$	19.383	0.450
	$\mathcal{GP}(X_1\boldsymbol{\beta}, K_{SE})$	14.797	-0.598
	$\mathcal{GP}(0, K_{Lin} + K_{SE})$	14.913	-0.939

A promising trend observed within Table 1 is that models using physics and data in combination outperformed both Morison's Equation and black-box data-based models. The residual modelling GP-NARX and combined kernel GP-NARX were the two best performing models, with similar NMSE scores and an improved MSLL on the combined kernel model. A potential reason for the worse MSLL of the residual model is that attempting to fit a GP to a residual rather than the complete signal greatly reduces the signal to noise ratio. This can make it challenging to pick out remaining structure within the signal, favouring long lengthscale, high noise variance model fits.

Inline with findings within the literature [20–23], Morison's Equation was found to perform very well considering its simplicity and outperformed the black-box GP-NARX. This was in part helped by the dataset, which was from a region of primarily unidirectional flow, a key assumption of Morison's Equation. The training, validation and test sets were all from drag-dominated flow regimes, meaning that parameters learned within the construction of the model were likely to be appropriate for the test conditions. When using Morison's Equation, it is important to consider the likely flow regime in which it may be implemented and adjust C_d and C_m accordingly.

Although achieving a moderate NMSE score, the black-box GP-NARX performed very poorly with regards to MSLL. This highlighted two issues: the importance of considering multiple performance metrics and unexpected behaviours that may arise from the use of black-box models, particularly highly flexible ones such as a GP-NARX. The cause of a poor MSLL here was the explosion of uncertainty intervals due to feedback of samples within the MC MPO prediction, a common phenomena within autoregressive model structures.

5.2 *Posterior contribution breakdown of model components*

Along with benefits in overall performance, an advantage of combining physics with data-driven techniques is the insight that can be provided by looking at the contribution of each model component. Whilst the separation of a mean function and residual model is already clear within the combined model structure, the individual kernel posterior contributions within a combined kernel require decomposition. Following [36] it is possible to derive the conditional predictive distribution for the contribution of a kernel K_i within a combined additive kernel of the form $K = \sum_{i=1}^{i=n} K_i$.

$$p(\boldsymbol{f_i^*}|\boldsymbol{f^*}, X^*, \boldsymbol{f}, X, \boldsymbol{\theta}) \sim \mathcal{N}(K_i^{*T}(\sum_{i=1}^{i=n} K_i)\boldsymbol{y}, K_i^{**} - K_i^{*T}(\sum_{i=1}^{i=n} K_i)K_i^{*}) \quad (9)$$

where $\boldsymbol{f_i^*}$ is the prediction contribution of kernel K_i within the combined prediction $\boldsymbol{f^*} = \sum_{i=1}^{i=n} \boldsymbol{f_i^*}$. The breakdown of the mean function and residual GP contributions is shown in Figure 2, with the breakdown of kernel component contributions of the combined kernel model shown in Figure 3.

In both cases, Morison's Equation is able to capture the majority of structure within the wave force via either the mean function or linear kernel component. Here, this provides interpretability in to the flow conditions, highlighting the presence of a drag dominated flow regime and primarily unidirectional flow. For alternative flow conditions, particularly those breaking the assumptions of Morison's Equation, the relative contributions of model components would likely be very different.

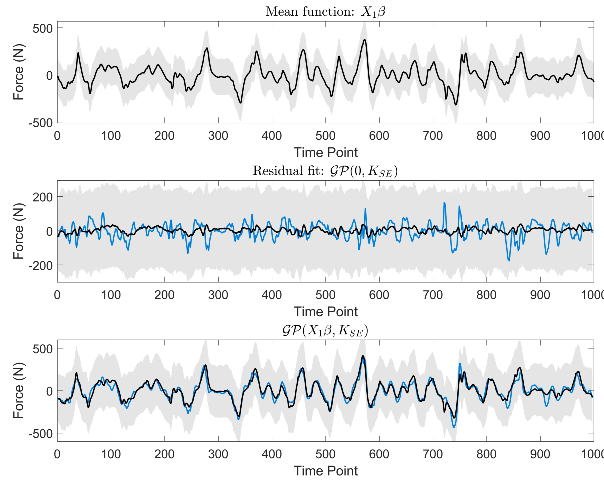

Figure 2: Residual modelling contribution breakdown of the linear mean function (top), residual GP-NARX with SE kernel (middle) and combined model (Bottom).

Due to the good fit of the mean function, the signal to noise ratio of the residual is very poor and the residual fit GP struggles to pick out the remaining structure. The middle plot of Figure 2 shows a long lengthscale fit with a large estimated noise. This contributes towards the larger variance and poorer MSLL of the combined residual model.

The consideration of individual component contributions within a model is of particular importance during extrapolation. Far from observed data, the data-based component of the model (or stationary kernel) will revert to its zero prior, with the overall model therefore dependant on the quality of either the mean function or physics-derived (nonstationary) kernel component. Independent of the overall model performance for a given testing dataset, ensuring that model components are capturing their intended processes is therefore important. The ability of kernel components to be learned simultaneously will help to reduce the 'biasing' phenomena of mean functions highlighted in Figure 1. Here we can see that the GP with the derived-kernels has captured each component very well, indicating that it is likely to perform well in extrapolation.

6 CONCLUSIONS

The use of Morison's Equation and a GP-NARX in combination was found to increase predictive performance over either technique used independently. Residual modelling was able to improve NMSE at the expense of MSLL, whilst the combined kernel was able to offer advantages in both NMSE and MSLL.

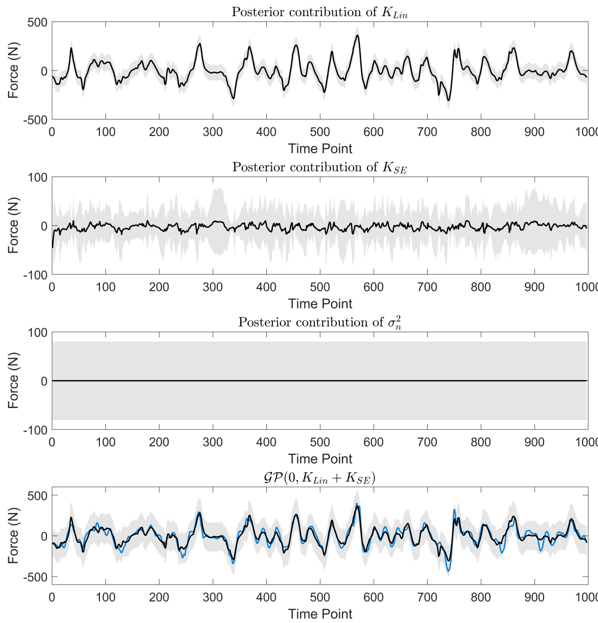

Figure 3: Combined kernel model contribution breakdown of the linear (Morison) kernel (top), SE GP-NARX (second), noise variance (third) and combined model (bottom).

Inline with the literature, Morison's Equation was found to achieve satisfactory performance when used in an appropriate flow regime for little computational cost. It provided a sensible start point for the development of combined physics and data-based models.

The decomposition of combined models was found to offer physical insight into the role of each model component. Here, due to the flow conditions, the majority of structure was captured via the physics-based component. Further work investigating how model component contributions vary over a range of flow conditions is planned to further explore how data-based learning may assist prediction in conditions where the performance of Morison's Equation would typically suffer.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support of Ramboll Energy, the University of Sheffield and the UK Engineering & Physical Sciences Research Council, through grant EP/S001565/1. We would like to thank Keith Worden for his support and for the provision of the CBT dataset.

REFERENCES

[1] Directorate of Energy Markets and Security. Renewables 2020 - analysis and forecast to 2025. Technical report, International Energy Agency, Renewable Energy Division, 2020.

[2] Paul Buchana and Patrick Mcsharry. Windstorm risk assessment for offshore wind farms in the north sea. *Wind Energy*, 22, 2019.

[3] Gerhard Ersdal and Erik Hörnlund. Assessment of offshore structures for life extension. In *Proceedings of the International Conference on Offshore Mechanics and Arctic Engineering - OMAE*, volume 5, 2008.

[4] Geoffrey Holmes, Pia Sartor, Stephen Reed, Paul Southern, Keith Worden, and Elizabeth J. Cross. Prediction of landing gear loads using machine learning techniques. *Structural Health Monitoring*, 15:568–582, 2016.

[5] T. J. Rogers, P Gardner, N Dervilis, K Worden, A. E. Maguire, E Papatheou, and E. J. Cross. Probabilistic modelling of wind turbine power curves with application of heteroscedastic Gaussian process regression. *Renewable Energy*, 148:1124–1136, 2020.

[6] Sikai Zhang, Timothy. J. Rogers, and Elizabeth J. Cross. *Gaussian Process Based Grey-Box Modelling for SHM of Structures Under Fluctuating Environmental Conditions*, pages 55–66. 2021.

[7] D. M. Hawkins. The problem of overfitting. *Journal of Chemical Information and Computer Sciences*, 44:1–12, 2004.

[8] B. Sohlberg and E. W. Jacobsen. Grey box modelling – branches and experiences. *IFAC Proceedings Volumes*, 41:11415 – 11420, 2008.

[9] T. J. Rogers, Geoffrey Holmes, E. J. Cross, and K. Worden. On a grey box modelling framework for nonlinear system identification. In *Special Topics in Structural Dynamics, Volume 6*, pages 167–178, 2017.

[10] K. Worden, R. J. Barthorpe, E. J. Cross, N. Dervilis, G. R. Holmes, G. Manson, and T. J. Rogers. On evolutionary system identification with applications to nonlinear benchmarks. *Mechanical Systems and Signal Processing*, 112:194 – 232, 2018.

[11] D. J. Pitchforth, T. J. Rogers, U. T. Tygesen, and E. J. Cross. Grey-box models for wave loading prediction. *Mechanical Systems and Signal Processing*, 159:107741, 2021.

[12] Prashanth Pillai, Anshul Kaushik, Shivanand Bhavikatti, Arjun Roy, and Virendra Kumar. A hybrid approach for fusing physics and data for failure prediction. *International Journal of Prognostics and Health Management*, pages 2153–2648, 2016.

[13] Alina Kloss, Stefan Schaal, and Jeannette Bohg. Combining learned and analytical models for predicting action effects. 10 2017.

[14] Houman Hanachi, Wennian Yu, Il Yong Kim, Jie Liu, and Chris Mechefske. Hybrid data-driven physics-based model fusion framework for tool wear prediction. *The International Journal of Advanced Manufacturing Technology*, 101:2861–2872, 2019.

[15] E. J. Cross, S. J. Gibson, M. R. Jones, D. J. Pitchforth, S. Zhang, and T. J. Rogers. Physics-informed machine learning for structural health monitoring. In *Structural Health Monitoring Based on Data Science Techniques*, pages 347–367. Springer, 2022.

[16] G. Najafian, R. G. Tickell, R. Burrows, and J. R. Bishop. The UK Christchurch Bay Compliant Cylinder Project: analysis and interpretation of Morison wave force and response data. *Applied Ocean Research*, 22:129 – 153, 2000.

[17] A. K. Agarwal and A. K. Jain. Dynamic behavior of offshore spar platforms under regular sea waves. *Ocean Engineering*, 30:487–516, 2003.

[18] J. Van Der Tempel, N. F. B. Diepeveen, W. E. De Vries, and D. Cerda Salzmann. 15 - Offshore environmental loads and wind turbine design: impact of wind, wave, currents and ice. In *Wind Energy Systems*, pages 463 – 478. Woodhead Publishing, 2011.

[19] J. R. Morison, Joseph W Johnson, and Samuel A. Schaaf. The force exerted by surface waves on piles. *Petroleum Transactions*, 189:149–157, 1950.

[20] Woodward-Clyde Consultants, Houston, Texas, Defense Technical Information Center. *Assessment of the Morison Equation*, 1980.

[21] D. J. Evans. Analysis of Wave Force Data. *Journal of Petroleum Technology*, 22(03):347–358, 1970.

[22] Graham R. Cook and Emil Simiu. Hydrodynamic forces on vertical cylinders and the lighthill correction. *Ocean Engineering*, 16(4):355–372, 1989.

[23] Turgut 'Sarp' Sarpkaya. *Wave Forces on Offshore Structures*. Cambridge University Press, 2010.

[24] Subrata K. Chakrabarti, Allan L. Wolbert, and William A. Tam. Wave forces on inclined tubes. *Coastal Engineering*, 1:149–165, 1977.

[25] V. Sundar, V. Vengatesan, G. Anandkumar, and A. Schlenkhoff. Hydrodynamic coefficients for inclined cylinders. *Ocean Engineering*, 25(4):277–294, 1998.

[26] G. Anandkumar, V. Sundar, K. U. Graw, and H. Kaldenhoff. Pressures and forces on inclined cylinders due to regular waves. *Ocean Engineering*, 22:747–759, 1995.

[27] James Lighthill. Fundamentals concerning wave loading on offshore structures. *Journal of Fluid Mechanics*, 173:667–681, 1986.

[28] T. Sarpkaya. Resistance in unsteady flow: Search for an in-line force model. *International Journal of Offshore and Polar Engineering*, 10:249–254, 12 2000.

[29] P. K. Stansby, K. Worden, G. R. Tomlinson, and S. A. Billings. Improved wave force classification using system identification. *Applied Ocean Research*, 14(2):107–118, 1992.

[30] K. Worden, P. K. Stansby, G. R. Tomlinson, and S. A. Billings. On wave force analysis using system identification. Technical report, University of Sheffield, 1992.

[31] Topi Paananen, Juho Piironen, Michael Andersen, and Aki Vehtari. Variable selection for gaussian processes via sensitivity analysis of the posterior predictive distribution. In *Proceedings of the 22nd International Conference on Artificial Intelligence and Statistics*, Proceedings of Machine Learning Research, 04 2019.

[32] Faizal Hafiz, Akshya Swain, Eduardo MAM Mendes, and Nitish Patel. Structure selection of polynomial NARX models using two dimensional (2d) particle swarms. In *2018 IEEE Congress on Evolutionary Computation (CEC)*, pages 1–8, 2018.

[33] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for Machine Learning*. The MIT Press, 2005.

[34] K Worden, William Becker, T. J. Rogers, and E. J. Cross. On the confidence bounds of Gaussian process NARX models and their higher-order frequency response functions. *Mechanical Systems and Signal Processing*, 104:188–223, 2018.

[35] T. J. Rogers. *Towards Bayesian System Identification: With Application to SHM of Offshore Structures*. PhD thesis, University of Sheffield, 2019.

[36] David Duvenaud. *Automatic model construction with Gaussian processes*. PhD thesis, University of Cambridge, 2014.